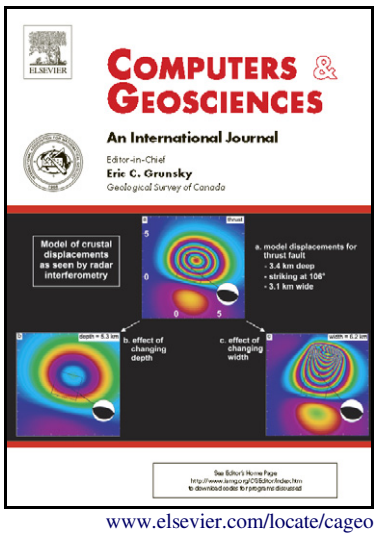
# Author's Accepted Manuscript




www.elsevier.com/locate/cageo

Doing Fieldwork on the Seafloor: Photogrammetric Techniques to yield 3D Visual Models from ROV Video

Tom Kwasnitschka, Thor H. Hansteen, Colin W. Devey, Steffen Kutterolf

PII: S0098-3004(12)00348-2
DOI: http://dx.doi.org/10.1016/j.cageo.2012.10.008
Reference: CAGEO3053

To appear in: *Computers & Geosciences*

Received date: 26 July 2012
Revised date: 10 October 2012
Accepted date: 11 October 2012

Cite this article as: Tom Kwasnitschka, Thor H. Hansteen, Colin W. Devey and Steffen Kutterolf, Doing Fieldwork on the Seafloor: Photogrammetric Techniques to yield 3D Visual Models from ROV Video, *Computers & Geosciences,* http://dx.doi.org/10.1016/j.cageo.2012.10.008

# Doing Fieldwork on the Seafloor:

# Photogrammetric Techniques to yield 3D Visual Models from ROV Video

**Tom Kwasnitschka**[1], Thor H. Hansteen[2], Colin W. Devey[3], and Steffen Kutterolf[4]

all at

GEOMAR | Helmholtz Centre for Ocean Research Kiel

Wischhofstr. 1-3, D-24148 Kiel, Germany

[1]tkwasnitschka@geomar.de, phone: +49 431 600-2136, fax: +49 431 600-2924 (Corr. Author)

[2]thansteen@geomar.de

[3]cdevey@geomar.de

[4]skutterolf@geomar.de

## Abstract

Remotely Operated Vehicles (ROVs) have proven to be highly effective in recovering well localized samples and observations from the seafloor. In the course of ROV deployments, however, huge amounts of video and photographic data are gathered which present tremendous potential for data mining. We present a new workflow based on industrial software to derive fundamental field geology information such as quantitative stratigraphy and tectonic structures from ROV-based photo and video material. We demonstrate proof of principle tests for this workflow on video data collected during dives with the ROV Kiel6000 on a new hot spot volcanic field that was recently identified southwest of the island of Santo Antão in the Cape Verdes. Our workflow allows us to derive three-dimensional models of outcrops facilitating quantitative measurements of joint orientation, bedding structure, grain size comparison and photo mosaicking within a georeferenced framework. The compiled data facilitate volcanological and tectonic interpretations from hand specimen to outcrop scales

based on the quantified optical data. The demonstrated procedure is readily replicable and opens up possibilities for post-cruise “virtual fieldwork” on the seafloor.

Keywords:
Physical Volcanology, Photogrammetry, Microbathymetry, Visualization, Grain Size Analysis, Underwater Navigation

## 1. Introduction

The scientific use of Remotely Operated Vehicles (ROVs) has traditionally concentrated on making visual observations, installing equipment precisely on the seafloor and the recovery of physical samples. While the samples can later be localized with high precision, the sampling decision is often based upon very limited visual information relayed through video cameras. Without comprehensive prior surveys of the study area on scales at which the ROV operates afterwards, researchers cannot be sure to be observing or sampling in the scientifically most relevant places.

During the RV Meteor 80/3 cruise to the Cape Verdes, we encountered a complex submarine cone field, called the Charles Darwin Volcanic Field (CDVF), located at a depth of 3500m off the Island of Santo Antão. Based on a ship-based bathymetric map gridded at 25m resolution (Figure 1) and on a scattered set of dredge samples, we identified targets for the detailed surveying of several structures using the ROV Kiel6000 vehicle. It quickly became clear that the observational protocols usually used during ROV operations were failing to capture much geological information (especially on the relationships between rock structures in 3D) which could greatly help interpretation of the samples post-cruise. This motivated the development of the photogrammetric mapping workflow described below, as available acoustic survey methods and video coverage did not allow adequate interpretation of the bedding structures and series of deposits which we observed. The detailed geological interpretation and resulting development of new concepts on eruptive mechanisms in the deep sea will be the subject of a companion paper (Kwasnitschka et al., subm.).

Our aim was to design a workflow that would allow examination of the outcrops indirectly after the dive, based on a three-dimensional digital model (e.g. Kreylos et al., 2006). This model should be accurate enough to allow quantitative measurements matching the precision

of terrestrial surveys. At the same time, the representation of geological features should be realistic and unobstructed by effects of the water in order to allow the development of discussable qualitative field impressions. The procedure should be stable, repeatable and reasonably fast while easy to use, allowing a first evaluation of new data while still at sea. The workflow should be open, follow industry standards, and offer interfaces to further third party data treatment procedures. We directed our attention towards existing industry-proven components that would allow integration on a user level and be widely available to other users without specialized programming knowledge (although a familiarity with the features of the software we used (which can be acquired from the software documentation and HowTo video tutorials) is assumed here).

## 2. Previous Work

Photogrammetry has previously been successfully applied to a variety of marine scientific disciplines, an overview of which is given by Johnson-Roberson et al. (2010). Our own work was inspired by recent applications in the field of aerial archeology (Verhoeven, 2011) with which we share the same general requirements of georeferenced visualization of terrain features on multiple scales. At the same time, photogrammetric methods have been developed to derive structural geology data from outcrops in time-critical situations such as tunneling and quarrying (Gaich et al., 2003). Our own previous use of such packages (Kwasnitschka, 2008) also informed the development of the current project.

The underwater application of photogrammetric techniques poses considerable problems since all camera equipment needs to be housed in some form of water-tight (and, in our case, pressure-resistant) housing. The air-glass-water interface introduced by this camera housing acts as an additional optical element - in the case of pressure housings, the optical geometry will also change in a possibly non-linear manner as its multiple components deform under

load. Various approaches have been taken to address these issues (e.g. Beall et al., 2010; Sedlazeck and Koch, 2011). Previous authors have investigated the optical aberrations induced by the air-water interface (e.g. Sedlazeck et al., 2009) providing critical constraints for positional reconstruction in the absence of reliable external navigation data.

Using established subaerial calibration methods, it is also possible to achieve locally well-defined reconstructions by using rigid stereoscopic camera rigs (Johnson-Roberson et al., 2010) or monoscopic approaches (Pizarro et al., 2009) and aligning clusters of such local models using vehicle navigation data. Much of the previous work has been made using images from Autonomous Underwater Vehicles (AUVs) rather than ROVs since they are generally able to cover larger areas due to their deployment scheme (Yoerger et al., 2007). For our scientific purposes, though, AUV have two major drawbacks: (a) they do not allow an immediate reaction to new discoveries and so often have objects of interest at the edges of images and (b) they generally view scenes from above, not the visualization angle most suited to seafloor geological interpretation, where dip angles are mostly low and so features of interest are best viewed by looking horizontally onto vertical walls.

## 3. Hardware

The ROV KIEL6000 is a 6000m-rated work class ROV built by Schilling Robotics, Inc.. Seven thrusters allow precise positioning; an automatic station-keep mode is available for stationary operation. The lighting is centered to the front within a cone of approximately 100°, featuring a total of 990 Watt. Table 1 and Figure 2 give an overview of the standard camera equipment available on the vehicle. In addition to this, an experimental stereo camera rig consisting of two Ocean Imaging Systems camera housings with Nikon D80 cameras and 20 mm optics was tested during the test deployments described here. The cameras were mounted parallel at a separation of 20 cm between the optical axes. All housings are equipped with flat glass ports. Only the stereo rig was successfully calibrated for optical aberrations

111 since the other cameras, mechanically and electrically fully integrated into the vehicle itself,
112 were unavailable for such a procedure.
113
114 The underwater navigation system installed on the vehicle consists of three elements. An RDI
115 Workhorse Navigator 1200 Doppler Velocity Logger (DVL) provides three-dimensional
116 differential motion tracking at sub-centimeter resolution. Absolute position determination is
117 supported by a Posidonia 6000 Ultra Short Baseline Logger (USBL) pinger installed on the
118 vehicle and used in combination with a USBL antenna mounted on the surface vessel (in this
119 case RV METEOR). A compass (Precision Navigation TCM2-50) provides information on
120 the vehicle heading. Depth and altitude readings were provided by a Seabird Electronics
121 FastCat SBE 49 CTD sensor, the DVL and by a Paroscientific Digiquartz pressure sensor.
122 Orientation data of the vehicle (heading, pitch and roll) were received from a Crossbow
123 VG700CA orientation sensor. Pan and tilt values of the navigation cameras were logged from
124 the Schilling rotary actuators.
125
126 **4. Data**
127 The outcrops which we used to test the 3D reconstruction workflow are situated on the
128 Charles Darwin Seamounts. They display abundant units with clearly visible bedding
129 structures cut by several generations of joints. Our test data was acquired on ROV dive 059
130 during Meteor cruise 80/3 which led across the northwestern flank of a volcanic crater
131 showing the morphological characteristics of a tuff ring (Figure 3). The ring rises about 200 m
132 above the surrounding terrain and its central depression measures 1 km across featuring steep
133 inner walls dropping step wise between 20 m and 50 m at angles between 35° and 90°. The
134 structure is composed of loose to strongly compacted volcaniclastics and lava flows, featuring
135 a sparse to locally moderate population of vagile and sessile epibenthic fauna.
136

The sample film clip (059-47) was recorded during a descent of the vehicle 28 m down a cliff of 65° inclination with sedimentary talus at the foot of the cliff marking the transition to the crater floor. Using the ROV´s onboard sonar, a roughly constant distance of 2.5 m to the wall could be maintained. Consequently, the width of the model ranges between 3.8 m and 8.5 m. The duration of the sequence is just over two minutes, resulting in a set of 125 source images after pre-processing. All but the last 7 images (where the image quality deteriorated with increasing distance to the wall) could be included in the reconstruction. This yielded a model of 200 000 polygons and a texture of 8192x8192 pixels, corresponding to an estimated geometrical resolution of 15 cm and a local maximum textural resolution of 2 cm. The textures of the uppermost and lowermost 25 % of the model in particular show signs of color absorption into bluish hues due to a slightly increased object distance. Since there was little variation in the horizontal viewing angle, some laterally facing portions, especially features facing right, remained occluded from the camera and could not be textured.

## 5. Data Pre-Processing

The individual streams of video and still imagery were synchronized in Adobe Premiere and merged into a master stream for reference. At this point it became clear that we would not be able to use the data from the experimental stereo camera system as it showed synchronization faults and had collected data only sporadically during the dive. Pre-selected video sequences covering outcrops were reformatted using Adobe After Effects. Elements obstructing the view such as visible parts of the vehicle and non-static objects were masked out from the video (as they are not part of the static seafloor which we wanted to reconstruct and so would have led to erroneous results during automated 3D reconstruction). No noise reduction filters were applied as we found that the resulting loss of detail caused reconstruction gaps on uniform surfaces. The clips were exported at a rate of one frame per second as still sequences with the mask embedded as an alpha channel. The complete workflow is illustrated in Figure 4.

The USBL (absolute position with relatively large errors) and DVL (relative position with smaller errors) navigational data underwent extensive correction as it formed the only frame of reference to which the quality of reconstructed models could be gauged. Using a semi-automatic Matlab routine, we generated a hybrid vehicle path, stabilizing the two data sources with respect to each other (Figure 3): The x and y components of both signals were filtered for system-inherent outliers, after which the short-wavelength component of the DVL was copied onto the long-wavelength component of the USBL signal. We take this to be a best-practice approach with the given data quality, with the drawback that, in passages where the DVL failed, the precision of the accepted position is diminished to an average of the USBL position and can only serve for rough georeferencing but not for quality control. Likewise, the vertical component of the DVL signal was corrected for drift effects and failed passages using a merged version of the CTD and Digiquartz pressure sensors.

## 6. Photogrammetric Reconstruction

A number of different processing approaches can be grouped under the term photogrammetry, employing heterogeneous image clusters, a single camera or stereo pairs. In order to mathematically reverse the projection of a camera and reconstruct three-dimensional information from two-dimensional images, the algorithms require the extrinsic camera parameters (essentially, camera position relative to the object) and intrinsic camera parameters (a description of the optical path of light through the lens onto the sensor). A truly accurate reconstruction can only be achieved if there is precise information on the intrinsic parameters and some minimal external information in the form of camera position and orientation or reference points on the object in order to establish absolute scale, orientation and position within a world coordinate system.

189 After tests using a variety of motion tracking packages (PFtrack5.0, Boujou; Condell et al.,
190 2006), open source and free bundle adjustment software (Bundler, Microsoft Photosynth;
191 Triggs et al., 2000, Snavely et al. 2010), we now rely on Agisoft Photoscan Professional, a
192 commercial suite for aerial photogrammetry. This software offers an integrated workflow
193 including a core of sturdy reconstruction and georeferencing tools along with sufficient means
194 to pre-process the input images and edit the finished models for further export. Verhoeven
195 (2010) gives a detailed overview of the software functionality and the basic process of (1)
196 model triangulation from image features matched across a cluster of overlapping images into
197 (2) a sparse estimate of the scene geometry from which (3) a dense model is derived, followed
198 by (4) surface modeling, and (5) texture generation (Figure 4). The range of features has since
199 been expanded for georeferencing based on landmarks visible on the model for which
200 coordinates have been determined by GPS. Since a high-resolution, georeferenced AUV map
201 was unavailable, we registered our models to the camera poses (i.e. position and orientation)
202 derived from the ROV motion record. The software first tries to georeference by a best-fit,
203 rigid seven-parameter transformation. In a second step, the reconstructed geometry and
204 camera poses are subjected to a nonlinear optimization process to the ROV vehicle track. For
205 all poses, the degree of misfit is reported, along with RMSE values.
206
207 Afterwards, the photographic texture can be exported as an orthophoto mosaic after re-
208 projecting the images from the camera poses back across the model surface. Various texture
209 blending rules can be chosen to treat overlapping areas. The most instructive results can be
210 acquired by always choosing the brightest available pixel from the set of overlapping pixels
211 which could be projected onto a given surface coordinate. This somewhat suppresses the
212 effects of light absorption through the water and brightness falloff towards the outer parts of
213 the cone of illumination. In the case of strongly varying object distances, more visually
214 pleasing results can be achieved by forming an average of overlapping pixels. A

mathematically correct treatment of this problem appears to be difficult as the distance of the objects to the light sources and the camera varies strongly compared to a case of flat sea bottom and a monotonous vehicle track, where such corrections have been successfully demonstrated (Sedlazeck et al. 2009; Johnson-Roberson et al., 2010).

The models can be exported to a variety of formats and geographic reference frames. The full-resolution model is exported in the Autodesk .3ds format containing the textured model and the camera positions to be used for further interpretation. The orthophoto mosaic is derived at a resolution of 5cm per pixel, which matches the average resolution of the input material. For immediate visualization purposes, the model resolution is diminished to 20 000 polygons, and the texture is downscaled from 8192x8192 pixel to 2048x2048 pixel to be compatible to real time viewing applications. Another .3ds version is saved, along with a georeferenced Collada model with a KML link, and a U3D file contained in a PDF.

## 7. Model Interpretation

The goal of our project was to derive quantitative data of geological structures beyond mere size and distance measurements. To achieve this, we edit the models in Autodesk 3dsMax, which counts among the standard tools of the CGI industry. One major drawback of the software is that geographic references are lost while scale and orientation are maintained. The 3D scenes build upon a cartesian coordinate system in meters but cannot deal with geographic positions due to limitations in computational precision facing large numbers (Mach and Petschek, 2007). A workaround is to define a reference point at the center of the working area and to arrange all data relative to that point.

To derive quantitative geological information from the model, we follow the same basic route (e.g. Jones et al., 2008): (1) Create an Autodesk 3dsMax helper object, (2) align and scale it to

match the geological feature to be measured, (3) read the respective property of the helper object. The additional benefit is a direct visualization of the measurements, which can later be refined to produce a visually informative illustration.

To measure the orientation of planar structures such as faults, joints or bedding planes, a planar Autodesk "section object" is placed on the model. The orientation of this "section object" is then adjusted until its intersection with the modeled seafloor matches that of the geological structure (Figure 5). The more surface relief there is on the model, the more accurate the measurement is. This procedure is not only an easy graphical way to determine the orientation of a planar geological structure; it also has some distinct advantages over, for example, using a compass in the field. Firstly it allows the orientation of features with ill-defined boundaries (such as banks of coarse gravel) to be accurately determined. Secondly it increases the sampling area for the orientation measurement, providing a more representative "average strike and dip" than a point measurement.

On a larger scale, a "master section" can be used to vertically slice the orientated "section objects" along with the outcrop model, to provide a proportionally accurate geological profile through the outcrop which can be directly imported into vector drawing programs (Figure 5a).

A grain size estimate can be derived from the models by creating appropriate spheres, ellipses or boxes around the clasts to be measured, and reading the size of the object along the respective axes of interest. Working with a multitude of individual measurements, objects can be batch renamed and assembled into groups while the arrangement of these groups into visibility layers allows order to be maintained. As the textural resolution is higher than the geometrical resolution, clasts can also be measured based on the texture alone, allowing work down to the centimeter scale.

Thanks to the powerful scripting interface of 3dsMax, we developed a suite of import and export scripts for camera poses, grain size data and orientation measurements converted to geologic notation, and to visualize statistical parameters of the reconstruction step. This feature also opens endless possibilities for future, case-sensitive data mining routines as well as further optimization of the models themselves. The scripts are available from the leader author on request.

## 8. Visualization

Next to the quantitative evaluation of the models, another main goal of our project is the enhancement of qualitative analysis by means of appropriate data visualization. Particularly ROV visual data is generally only made available to users in the geologically irrelevant temporal dimension (time stamped rather than georeferenced) meaning that much of the information it contains is difficult to access retrospectively. We wanted to use the 3D models to transpose the ROV visual data set into a geographical frame of reference and so provide access to the video information via its position.

Once more, 3dsMax serves as a powerful editor to prepare the models for use in real time visualization software. But in addition it is also a visualization tool in its own right. Several outcrop models can be loaded at once, and upon re-establishing their relative positions, the correlation of geologic features between adjacent outcrops is as natural as in the field on land. The software allows appropriate visualization geometry to be constructed to illustrate the findings. Camera objects represent the poses in the scene. Since these originate from a film sequence and the ROV followed a track, the path of the ROV across the outcrop can be animated and navigated using a time line.

The pre-processed ROV track record can be visually compared against the reconstructed ROV path, along with digital elevation models of the local bathymetry, still and video footage, and

any other 2D or 3D content that can be referenced within the scene's coordinate system (e.g. Kwasnitschka, 2008). Thus, this is the only software applied in our study allowing the simultaneous visualization of all our data sets in four dimensions. At the same time, it is the only fully featured application able to actively manipulate the models in order to create new data products.

We use a number of other visualization platforms that mostly allow the passive interaction with the data:

- A birds-eye view computer animation of the reconstructed scene is rendered by 3dsMax and added to the original Adobe Premiere video composition allowing the comparison of the video material to a four dimensional, animated map.
- Various applications focus on the geospatial aspect of our data, first and foremost an ArcGIS project including the bathymetric map, the ROV track, the photo mosaics of the reconstructed outcrops, event marks such as sample locations and the final geologic map layers. Some limited control over the temporal coordinates of data is available, too.
- The bathymetric post processing and visualization software Fledermaus is capable of displaying most of the GIS layers within a four dimensional space. The simulation is based around a bathymetrical height field and the animated ROV track. The mode of visualization is passive as the software is designed primarily for the dissemination of bathymetry data, which has to be pre-processed. A major drawback to date is the restriction of import interfaces to support merely untextured models or monochrome point clouds providing a relatively inferior representation of the outcrop reconstruction effort.
- Virtual Globes such as Google Earth and World Wide Telescope digest distribution formats (photo mosaics, .3ds and Collada models using accompanying KML files)

directly exported from Photoscan. By the nature of the applications, navigation on a scale of meters is challenging.

- An even easier way to examine the unedited reconstructions with low-level tools is the U3D format contained in PDF format, which allows passive interaction with the model through Adobe Acrobat or Reader on any operating system. Quantitative measurements of sizes, angles and even orientation can be made.

KML samples of the data discussed in this article can be downloaded as online supplements.

## 9. Consideration of errors

In the absence of hard constraints on size and orientation of seafloor features provided by artificial gauges (e.g. parallel laser beams), evaluation of errors in the reconstructed geometry had to rely on indirect methods and assessment calculations in which we vary influences of imperfect navigation and optical distortion.

We assume the positional uncertainty within the navigational data to exceed the drift of the geometrical reconstruction. We infer this from a comparison of the reconstructed orientation data with the orientation data derived from the on-board sensors (certified to be accurate within one degree). The in-situ optical distortion parameters could not be constrained rigorously. Nevertheless, examining the contribution of different parameters to the overall result can elucidate the robustness of the reconstruction method. In all of the cases outlined below, nonlinear optimization was omitted in order to reveal the differences between acoustic navigation and optical measurements.

For dive 059, with a total distance traveled of 3354 m at an average of 0.076 m/s during 12.2 h (Figure 3), the expected drift of the DVL sensor based on manufacturer´s specification lies at +-95 m. Nevertheless, we found that, relative to USBL fixes (which have lower precision (quoted by the manufacturer as 1% of the operational water depth) but should not be subject

345 to time-dependent drift) our instrument deviated by 458 m (24% of the true distance between
346 start and end points) even after cleaning of the DVL record to remove obviously erroneous
347 episodes (e.g., apparent movement of vehicle above its maximum velocity or movement when
348 the video showed the vehicle was stationary). The USBL on RV Meteor had been calibrated
349 at depths comparable to our deployment relative to a fixed seafloor beacon on a preceding
350 cruise and been shown to have positional RMS error values of 13.7 m (x), 13.3 m (y) and 6.5
351 m (z) relative to ship's axis.
352
353 In our example model, the overall, timestamp corrected positional offset of the reconstruction
354 against the track is 0.143 m as laid out in Table 2. Figure 6 shows a graphical representation
355 of such offsets based on motion tracks and illustrates that, despite considerable uncertainties
356 in the overall positioning, the local fit among acoustical navigation and optical derivation of
357 camera poses can be very good. The concentration of positional misfit (Figure 6a) around
358 turning points throughout the sequence suggests that high frequency movement has been
359 suppressed during the track synthesis or that the DVL sensor occasionally exaggerated the
360 amount of vehicle motion in cases of fast movement. Meanwhile, the resulting differences in
361 camera orientation stay below 0.1°. The overall accuracy of the reconstruction can be further
362 constrained by a comparison of the camera orientation data to the original track, which yields
363 a cumulative average deviation of 5.2° (Figure 6c, Table 2).
364
365 A vital prerequisite for successful georeferencing is the availability of the high frequency
366 DVL signal. Referencing the data only on the long wavelength component of the USBL
367 signal increases the positional misalignment and suggests a temporal misalignment of 3
368 seconds between ROV track and reconstructed track towards a new positional alignment
369 optimum resulting in an ever increasing orientation imprecision (Table 2). In order to
370 illustrate the importance of correct timing, we compare the alignment at minus six seconds to

371 the alignment at plus six seconds, resulting in comparable deviations in orientation but
372 leading to a considerable degradation in positional accuracy. It should be noted though that
373 the degree of misalignment also depends on the length and complexity of the tracks to be
374 fitted. It is advisable to work with as large a model as possible in order to suppress local
375 disturbances.
376
377 The temporal resolution required is governed by the amount of overlap between the images
378 and thus by the speed of the ROV over ground. We achieve good results for a one-second
379 interval at vehicle speeds around 0.25-0.5 m/s, although a two or three-second interval often
380 still provides sufficient information. To investigate effects of varying data coverage on
381 reconstruction quality and to define a minimum shooting interval, we attempted
382 reconstruction from a subset of photographs chosen at ever-increasing intervals. For our
383 setup, we find a sampling frequency of 0.5 Hz (one image every 2 seconds) or higher to be
384 sufficient. The models created at one and two-second sampling intervals are essentially
385 identical in their extent, resolution and orientation (Table 2). Longer intervals create coverage
386 gaps that cause the reconstruction to halt. Since the algorithm does not look for unlinked
387 clusters beyond the first one it finds, the remaining images must be identified and a
388 reconstruction may be attempted in a new project. Such unlinked snippets of the entire
389 outcrop must then be individually georeferenced, introducing additional errors.
390
391 To illustrate the effect of an erroneous camera calibration we removed the calibration data for
392 the nonlinear lens distortion parameters K1-K3 (Brown, 1966). This resulted in a warping of
393 the scene along the x-axis expressed by a propagating deviation of the camera tilt, the location
394 of the point of view and a corresponding warping of the model around the x-axis and vertical
395 stretching. Instead of using the usual rigid alignment to the known vehicle track, we
396 superimposed the initial (i.e. the topmost) camera poses of the "correct" (i.e. with camera

397 calibration) and warped reconstructions to observe the propagation of drift throughout the
398 model. Deviations developed as given in Table 2 resulting in an offset of 7.6 m in position (-
399 4.6 m in x, 5.1 m in y and -3.2 m in z) for a feature at the bottom of the model. The largest
400 offset is experienced for any orientation measurements of faults and bedding planes towards
401 the lower end of the model, which are effectively rendered useless as illustrated in Figure 7.
402 Once the usual rigid body alignment to the vehicle track was carried out, the misfit was
403 distributed throughout the model but should still be regarded as unacceptable (Table 2).
404 Moreover, a strong deviation in scale becomes apparent. This experiment underlines the vital
405 importance of proper calibration of intrinsic camera parameters.
406
407 **10. Lessons Learned**
408 Having developed a workflow based on already existing data, we present here a number of
409 observations and best practice rules to be applied during data acquisition which, with minimal
410 additional effort when preparing and carrying out a ROV dive, can greatly improve the
411 quality and quantity of subsequent reconstructions.
412
413 1. First and foremost, the quality of reconstructed models depends critically on the
414 quality of the camera technology used to acquire the data. Using a high-resolution
415 sensor with a low noise level is the critical factor, as poor image quality significantly
416 contributes to erroneous matching of images. The optical system should be as simple
417 as possible, utilizing a lens with a very wide depth of focus or fixed focus balanced
418 with a small relative aperture.
419 2. Although an optically corrected "dome port" window for the camera pressure housing
420 is ideal, a minimal requirement is that corrections can be made for optical aberrations
421 due to the air/glass/water interface, requiring in-situ calibration of the camera image
422 using appropriate test objects (e.g., a checkerboard).

3. While the wish to invest in multifunctional deep-diving camera equipment to satisfy a range of user requirements is understandable, it should be realized that the use of zoom optics during photogrammetry is highly detrimental and can only be accommodated if focal length can be logged accurately, and the concomitant changes in optical aberration calculated, for every picture taken. Additionally the reconstruction software must be able to incorporate these data into its model-generation process; otherwise zoom will be interpreted as motion closer to the object, rendering track alignment impossible.
4. Photogrammetry is a computationally very intensive process. Our project was processed on a workstation equipped with a twin Intel Xeon Processor, 12 GB RAM and one ATI Crossfire 4800 Series graphics card (which is used with the GPU features of Photoscan), which matched our HD video source in required performance. A crucial bottleneck was the amount of RAM available. With this hardware, a 200 000 polygon reconstruction based on one-second samples of two minutes of HD video required just over four hours to process. The camera should be held orthogonal to the objects, angles larger than 45° may lead to a failure in reconstruction or to inaccuracies. Rapid movements (which lead either to blurring of the images or inaccuracies in the vehicle´s position and attitude determinations, possibly as a result of timing errors) should be avoided. An ideal platform for photogrammetric survey in the deep sea would move on a continuous gridded track ("mowing the lawn"). Tracks should cross each other frequently or ideally run parallel with considerable overlap.

Finally, it should be noted that our selection of software to pre-process and interpret the data is to be regarded as preliminary. Originating from the entertainment industry, many programs are barely able to cope with the amount of data generated at sea. The long duration of ROV dives requires workarounds in video editing, and 3D animation packages also require a

448 rescaling of geographic coordinates and time (e.g. minutes expressed as seconds to fit a
449 limited timeline).
450
451 **11. Conclusion**
452 We present a review of available software and a practical workflow to virtually replicate
453 morphological, geological and biological features of the seafloor accurately enough to
454 conduct scientific studies. Our aim was to create synthetic model visualizations of the seafloor
455 to provide more geologically useful and quantitative information than individual video
456 frames. Although we recognize that technical improvements can be made in order to further
457 substantiate and calibrate the interpretations inferred from our surveying method, we have
458 demonstrated that a georeferenced reconstruction of the features imaged during a video
459 transect is (a) possible and (b) yields quantitative geological information not otherwise
460 accessible. The example presented shows the potential of quantitative geoscientific
461 measurements and fieldwork on the deep ocean floor provided there is sufficient coverage by
462 adequate surveying. Similar benefits can be expected for biological habitat mapping as well as
463 for a wide variety of industrial seabed monitoring applications. At the same time, our method
464 is capable of processing archived data meeting minimal standards in order to re-evaluate
465 remote targets or to monitor changes at sites of high temporal variability such as hydrothermal
466 vents. In addition, the method provides a means to re-cast ROV video data in a geographical
467 rather than a time-line reference frame.
468
469
470 **Acknowledgements**
471 We would like to thank the Captain and crew of RV Meteor for their excellent assistance at
472 sea, and the crew of ROV Kiel6000 for their support during the development of such a
473 specific application, and for clarifying many engineering details. T. Kwasnitschka would like

to thank Anne Jordt for an introduction to underwater photogrammetry and many valuable comments. This work was carried out within the Jeddah Transect project (www.jeddah-transect.org).

**References:**

Beall, C., Lawrence, B.J., Ila, V., Dellaert, F., 2010. 3D reconstruction of underwater structures. In: Proceedings of Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference, pp.4418-4423, doi: 10.1109/IROS.2010.5649213

Brown, D.C., 1966. Decentering distortion of lenses, Photogrammetric Engineering. 32 (3), pp. 444–462.

Condell, J., Moore, G., Moore, J., 2006. Software and Methods for Motion Capture and Tracking in Animation. In: Arabnia, H. R. (Ed.), Proceedings of 2006 International Conference on Computer Graphics & Virtual Reality (CGVR'06), Las Vegas, pp. 3-9.

Gaich, A., Fasching, A., Fuchs, R., Schubert, W., 2003. Structural rock mass parameters recorded by a computer vision system. In: Culligan, P.J., Einstein, H.H., Whittle, A.J. (Eds.), Proceedings of Soil and Rock America 2003, 12th Panamerican Conference on Soil Mechanics and Geotechnical Engineering, Cambridge, MA, pp. 87-94.

498 Johnson-Roberson, M., Pizarro, O., Williams, S.B., Mahon, I., 2010. Generation and
499 visualization of large-scale three-dimensional reconstructions from underwater robotic
500 surveys. Journal of Field Robotics, 27 (1), pp. 21-51.
501
502 Jones, A.M., Cantin, N.E., Berkelmans, R., Sinclair, B., Negri, A.P., 2008. A 3D modeling
503 method to calculate the surface areas of coral branches. Coral Reefs, 27, pp. 521-526,
504 doi: 10.1007/s00338-008-0354-y.
505
506 Kreylos, O., Bernardin, T., Billen, M.I., Cowgill, E.S., Gold, R.D., Hamann, B., Jadamec, M.,
507 Kellogg, L.H., Staadt, O.G., Sumner, D.Y., 2006. Enabling scientific workflows in
508 virtual reality. In: Proceedings of the 2006 ACM international conference on Virtual
509 reality continuum and its applications, ACM, New York, pp. 155-162, doi:
510 10.1145/1128923.1128948.
511
512 Kwasnitschka, T., 2008, Volcanic and Tectonic Development of the Ilopango Caldera, El Sal-
513 vador: stratigraphic correlation and visualization of emplacement. Unpublished
514 Diploma Thesis, Christan-Albrechts-Universität zu Kiel, Kiel, 218pp.
515
516 Kwasnitschka, T., Hansteen, T., Devey, C.W., Kutterolf, S., Freundt, A., 2012. Stratigraphy
517 and Volcanotectonics of Deep-Sea Pyroclastic Deposits at Charles Darwin Volcanic
518 Field, Cape Verdes. Subm. Geochemistry, Geophysics, Geosystems.
519
520 Mach, R., Petschek, P., 2007. Visualization of Digital Terrain and Landscape Data, Springer,
521 New York, 374pp.
522

523 Pizarro, O., Eustice, R., 2009. Large area 3-d reconstructions from underwater optical
524 surveys, IEEE Journal of Oceanic Engineering, 34 (2), pp. 150-169.
525
526 Sedlazeck, A., Köser, K., Koch, R., 2009. 3D Reconstruction Based on Underwater Video
527 from ROV Kiel 6000 Considering Underwater Imaging Conditions. In: Proceedings of
528 IEEE OCEANS 2009, Bremen, pp. 1-10.
529
530 Sedlazeck, A., Koch, R., 2011. Calibration of Housing Parameters for Underwater Stereo-
531 Camera Rigs. In: Hoey, J., McKenna, S., Trucco, E. (Eds.), Proceedings of the British
532 Machine Vision Conference, BMVA Press, pp.118.1-118.11.
533
534 Snavely, N., Simon, I., Goesele, M., Szeliski, R., Seitz, S.M., 2010. Scene reconstruction and
535 visualization from community photo collections. Proceedings of the IEEE, 98 (8), pp.
536 1370-1390.
537
538 Triggs, B., McLauchlan, P.F., Hartley, R.I., Fitzgibbon, A.W., 2000. Bundle Adjustment - A
539 Modern Synthesis. In: Vision Algorithms: Theory and Practice, Lecture Notes in
540 Computer Science, 1883, Springer, Berlin / Heidelberg, pp. 153-177.
541
542 Verhoeven, G., 2011. Taking computer vision aloft – archaeological three-dimensional
543 reconstructions from aerial photographs with Photoscan. Archaeological Prospection
544 18, pp. 67-73.
545
546 Yoerger, D., Bradley, A., Jakuba, M., German, C.R., Shank, T., Tivey M., 2007. Autonomous
547 and remotely operated vehicle technology for hydrothermal vent discovery, exploration,
548 and sampling. Oceanography, 20 (1), pp. 152-161.

549 **Figure Captions:**

550 Figure 1:

551 Location of the working area. a) Asterisk marks the location of CDVF on the southwestern
552 slope of the Island of Santo Antão. b) Box marks the working area of ROV dive 059 at
553 Tambor Cone, part of the central group of CDVF (compare Fig. 4 for close up view).

554

555 Figure 2:

556 Cameras, positional sensors and lighting of ROV Kiel6000 seen from a) front and b) right
557 views. Relevant lighting equipment is labeled as c) Flash guns for the stereo system (OE11-
558 242), d) Halogen lamps (Deep Multi-SeaLite) e) HMI lamps (SeaArc 2), f) HID lamps
559 (SeaArc 5000). The stereo camera rig was mounted on a retractable sled and is shown in
560 retracted position, while during operation the lenses were positioned over the front edge of the
561 vehicle. The DVL and orientation sensors are obstructed. Approximate positions of the
562 cameras and sensors are given relative to g) lower front port corner of vehicle. Photos
563 courtesy H. Huusmann and N. Augustin.

564

565 Figure 3:

566 The Vehicle track of dive 059 used for alignment of the reconstructed geometry data consists
567 of the merged DVL bottom velocity and USBL data substituted by just the smoothed USBL
568 data where DVL data was unavailable (a). The grey arrow marks the accumulated offset
569 (458m) between the raw USBL and uncorrected DVL signal by the end of the dive.

570

571 Figure 4:

572 Workflow scheme developed during this study employs a number of commercial software
573 products which are linked using their own data interfaces as well as custom import and export
574 scripts for quantitative interpretation.

Figure 5:

a) Vertical profile, geometric features and stratigraphic units of outcrop 059-47 derived from b) the outcrop model. Lateral and vertical extent match in dimensions. The white dashed line marks the location of the profile. Linears on the model mark the intersection of the rock face and the joints and bedding planes. Note the white semitransparent example joint plane with indicators for strike (line) and dip (arrow) directions. c) shows a frame of the original video sequence at its approximate position in the outcrop (grey arrow).

Figure 6:

Visualization of the deviation of reconstruction values against the navigational data. a) Spatial deviation along the track path is plotted as a color coded spherical marker with the radius of the resulting error around the assumed position. b) Spatial deviation of the individual axes along the track including the resulting error vector. c) Orientation deviation along the three axes.

Figure 7:

Warping effects due to missing lens distortion parameters (a) superimposed on the correct reconstruction (b). Both models have been aligned at the first camera pose (c), where deviations in the model geometry and position are already apparent. The largest dislocation (gray arrow) in position and camera angle is found between the last images, (d) showing the warped path and (e) the correct path, deviating 29° in pitch, 8° in roll and 1.8° in heading. Crosses mark the location of a corresponding feature referenced in the text. Measurements of a corresponding bedding plane (white planes) indicate a strong deviation in strike (67°, lines) and dip (12°, arrows). The light transparent model (f) and camera planes (g) illustrate the

model, which has been aligned to the track coordinates, resulting in positioning and also scaling errors. The white grid represents the true horizontal plane.

**Tables:**

**Table 1 Relevant camera equipment installed on ROV Kiel6000. (Only HD and stereo cameras were used for reconstruction.)**

| | Brand | Recording format | FoV | Type of mount | Operation mode |
|---|---|---|---|---|---|
| **Navigation Cameras** | Kongsberg OE14-366 MKII | PAL | 63° | pan/tilt, logged | continuous |
| **HD Cameras** | Kongsberg OE14-500 | HDV, 1080i | 50° | tilt, unlogged | on demand |
| **Rear Navigation Camera** | Oktopus 6000 | PAL | 75° | fixed | not recorded |
| **Stereo Camera** | Ocean Imaging Systems | 10.2MP RAW | 57° | fixed | 4 sec interval on demand |
| **Still Camera** | Kongsberg OE14-208 | 5MP JPEG | 62° | pan/tilt, logged | on demand |

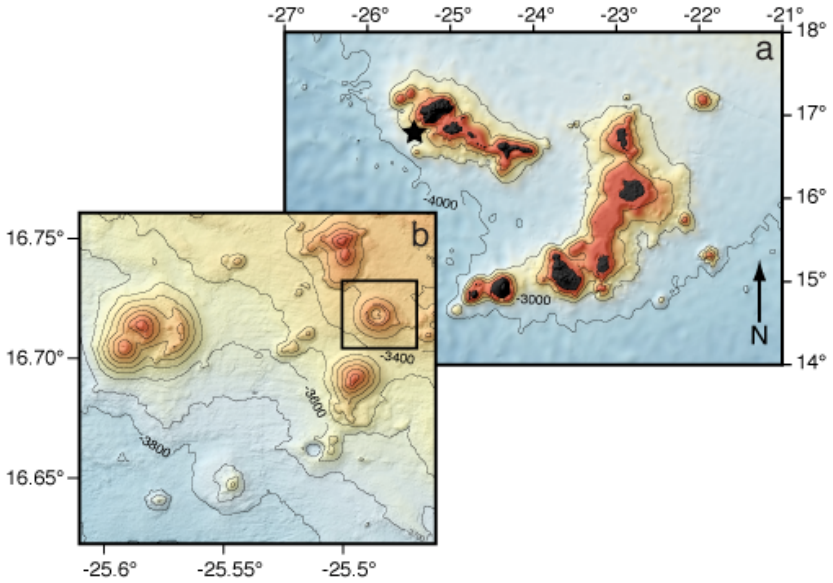
**Table 2 Top to bottom: Pose deviation of the best reconstruction attempt compared to effects of improper filtering, timing, capture interval, missing calibration aligned to the first pose (1) and the attempt to fit the uncalibrated model to the track record (2). Values are averages or root mean square errors of the entire series with respect to the telemetry data.**

| | **Orientation Differences (Average in °)** | | | **Positional Differences (RMSE in m)** | | | |
|---|---|---|---|---|---|---|---|
| | pitch | roll | heading | err x | err y | err z | net |
| **proper** | 3.2 | 0.0 | 2.0 | 0.10 | 0.06 | 0.08 | 0.14 |
| **USBL smooth** | -22.3 | 13.9 | 50.9 | 0.24 | 0.63 | 0.23 | 0.71 |
| **USBL -3 sec.** | -24.5 | 14.0 | 57.1 | 0.17 | 0.63 | 0.23 | 0.69 |
| **six sec. early** | 1.2 | -0.9 | 1.8 | 0.23 | 0.21 | 0.38 | 0.49 |
| **six sec. late** | 4.7 | -0.4 | 3.8 | 0.24 | 0.17 | 0.27 | 0.40 |
| **0.5 Hz interval** | 3.1 | -0.2 | 2.3 | 0.10 | 0.06 | 0.09 | 0.14 |
| **missing calib. (1)** | 20.0 | 2.9 | 1.0 | 0.54 | 1.41 | 2.10 | 2.58 |
| **missing calib. (2)** | 1.9 | -4.4 | 10.7 | 0.55 | 0.24 | 0.51 | 0.79 |

621

622 Figures:

623 Figure 1:



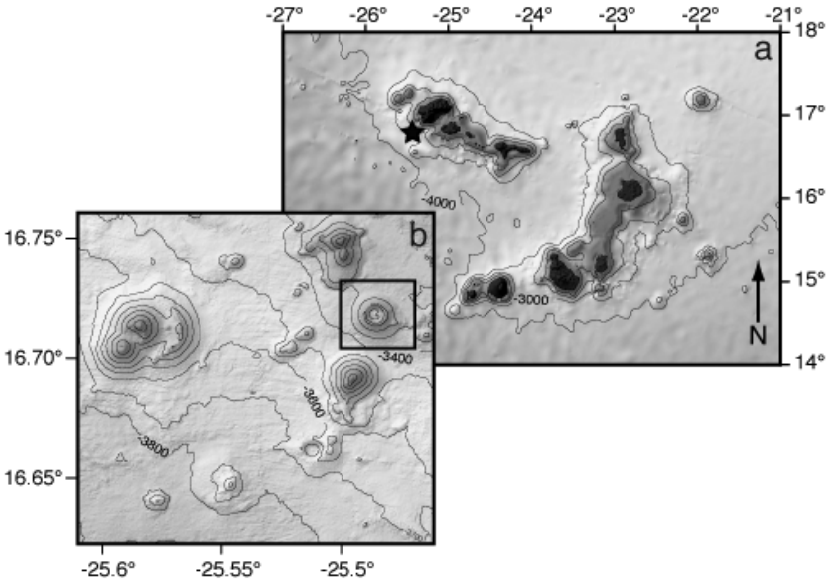

624

625

626

627 Figure 1, black and white:

628



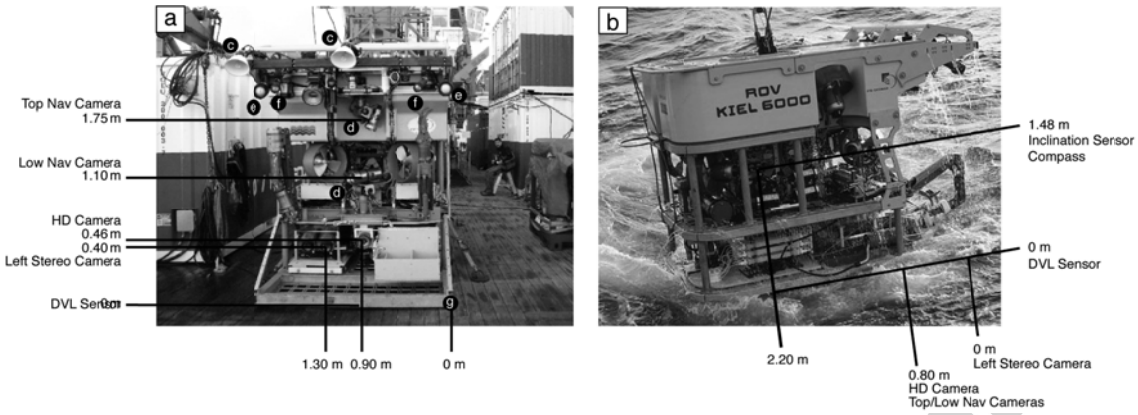

629

630

631

632 Figure 2, black and white:

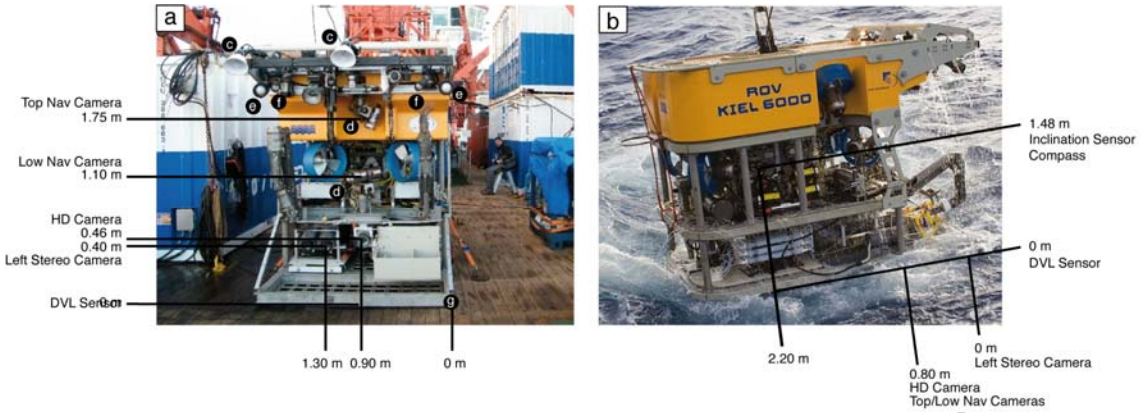
633




634

635

636 Figure 2, color:



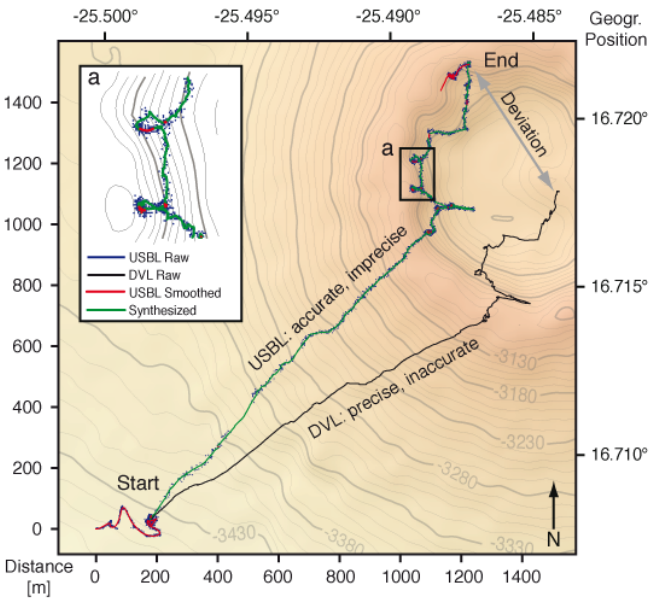

637

638

639

640

641
642 Figure 3, color:

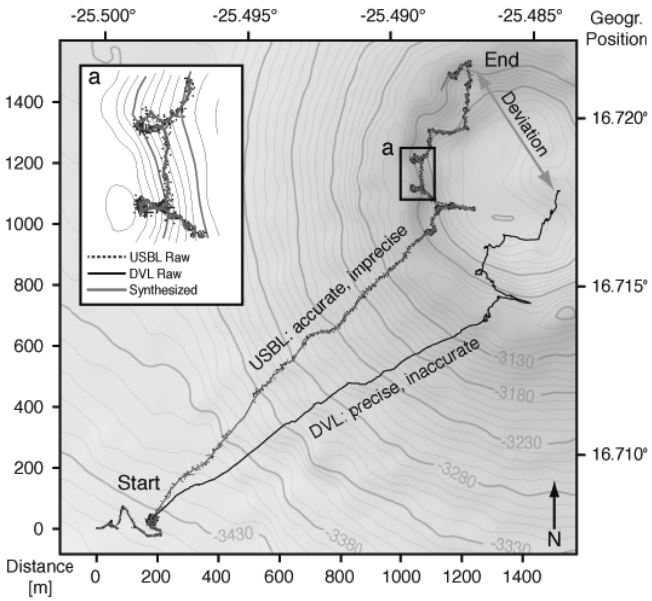



643
644

645

646

647 Figure 3, black and white:

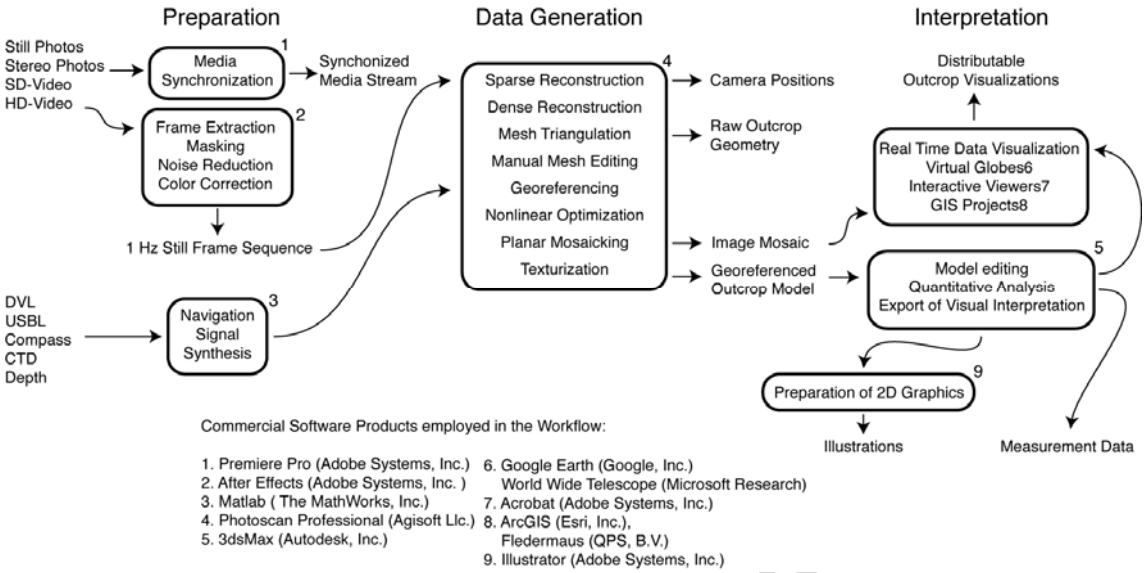



648

649

650

651 Figure 4:

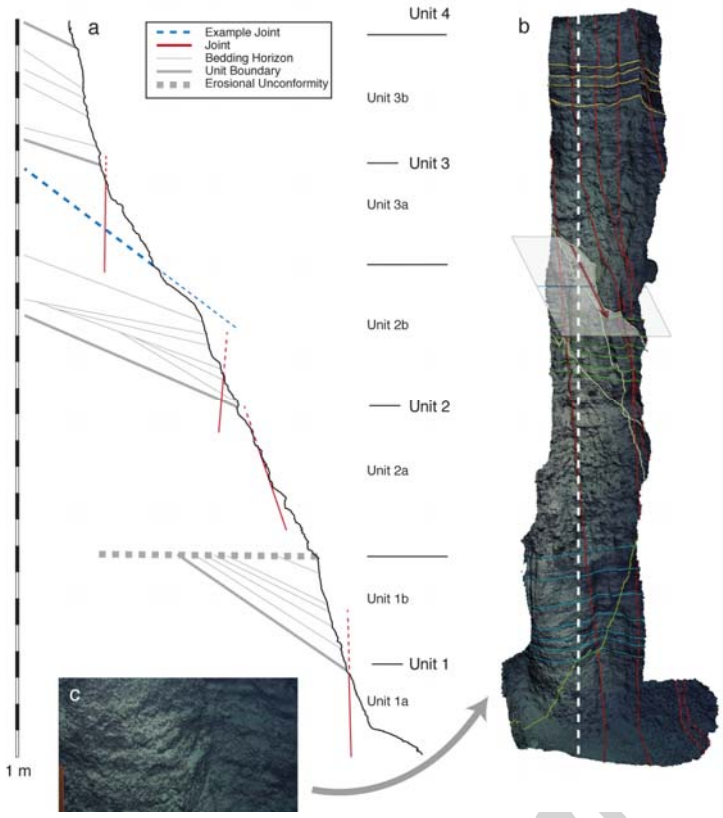



652

653

654

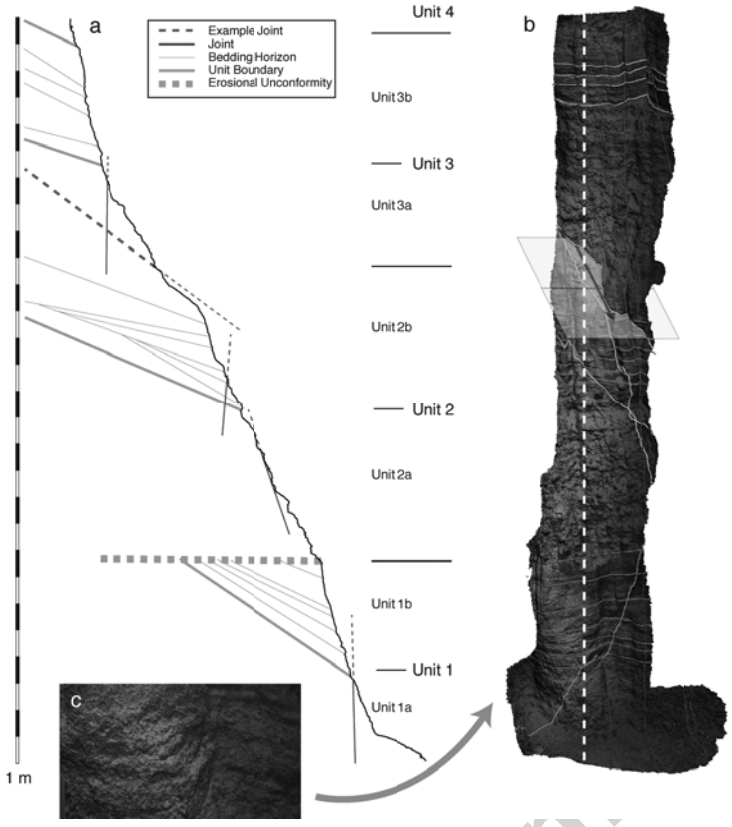
655 Figure 5, color:




656

657

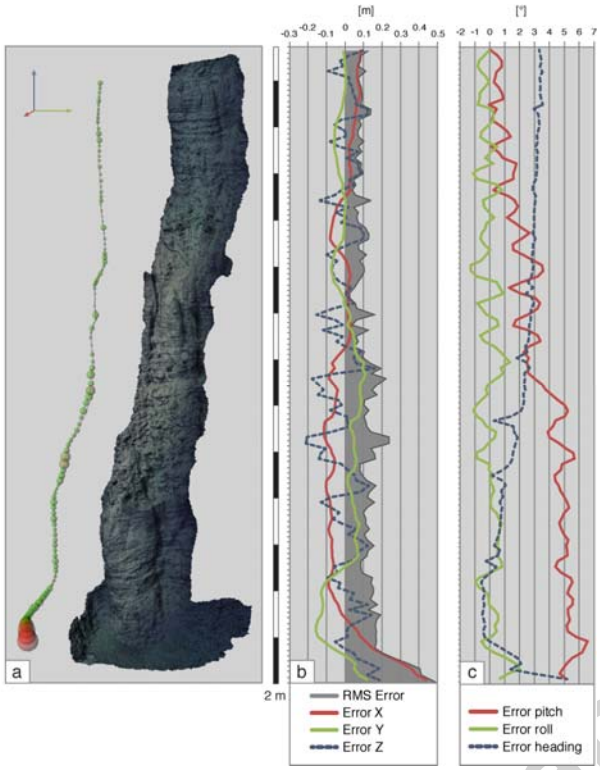
658 Figure 5, black and white:




659

660

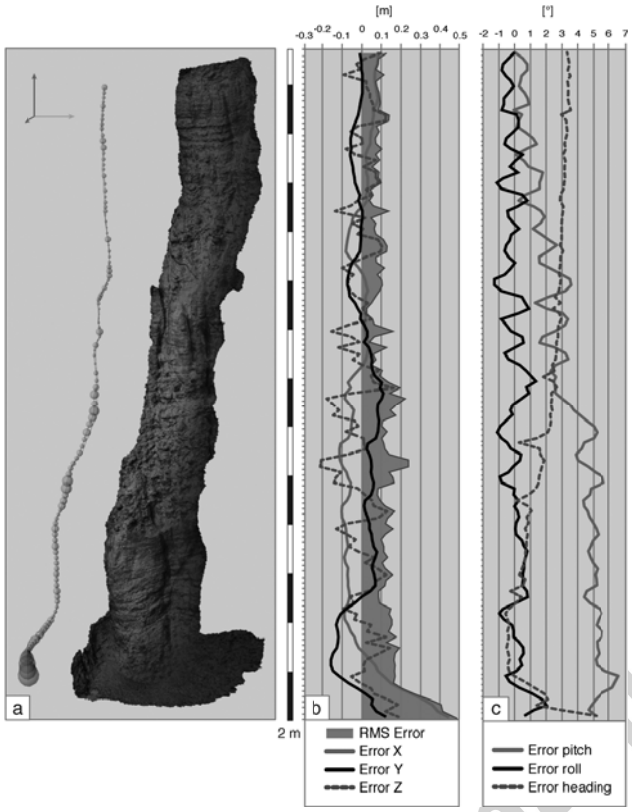
661 Figure 6, color:




662

663

664 Figure 6, black and white:



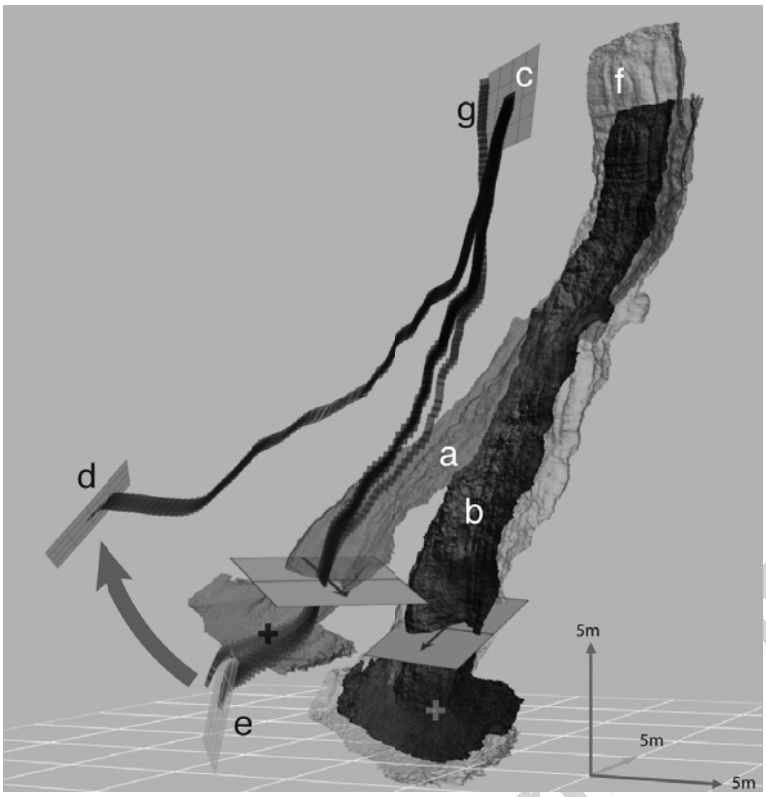

665

666

667

668

669 Figure 7, black and white:

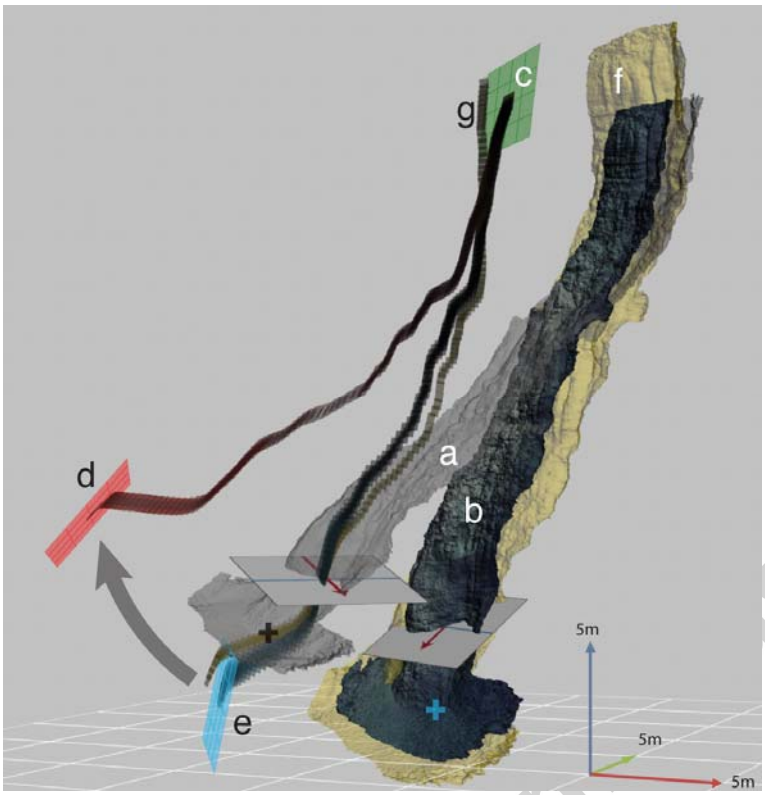



670

671

672

673 Figure 7, color:




674

675

Highlights:

- A new technology for deep-sea micro scale mapping is demonstrated.
- Photogrammetry based on ROV video yields 3D models.
- Quantitative data extraction yields geoscientific insights.
- The workflow is readily replicable and based on industrial software.